

| Minimum<br>signal<br>peptide score | false positive<br>rate | false<br>negative rate | proba(0.1) | proba(0.2) |
|------------------------------------|------------------------|------------------------|------------|------------|
| 3,5                                | 0,121                  | 0,036                  | 0,467      | 0,664      |
| 4                                  | 0,096                  | 0,06                   | 0,519      | 0,708      |
| 4,5                                | 0,078                  | 0,079                  | 0,565      | 0,745      |
| 5                                  | 0,062                  | 0,098                  | 0,616      | 0,782      |
| 5,5                                | 0,05                   | 0,127                  | 0,659      | 0,813      |
| 6                                  | 0,04                   | 0,163                  | 0,694      | 0,836      |
| 6,5                                | 0,033                  | 0,202                  | 0,725      | 0,855      |
| 7                                  | 0,025                  | 0,248                  | 0,763      | 0,878      |
| 7,5                                | 0,021                  | 0,304                  | 0,78       | 0,889      |
| 8                                  | 0,015                  | 0,368                  | 0,816      | 0,909      |
| 8,5                                | 0,012                  | 0,418                  | 0,836      | 0,92       |
| 9                                  | 0,009                  | 0,512                  | 0,856      | 0,93       |
| 9,5                                | 0,007                  | 0,581                  | 0,863      | 0,934      |
| 10                                 | 0,006                  | 0,678                  | 0,835      | 0,918      |

FIGURE 2

## Score curves

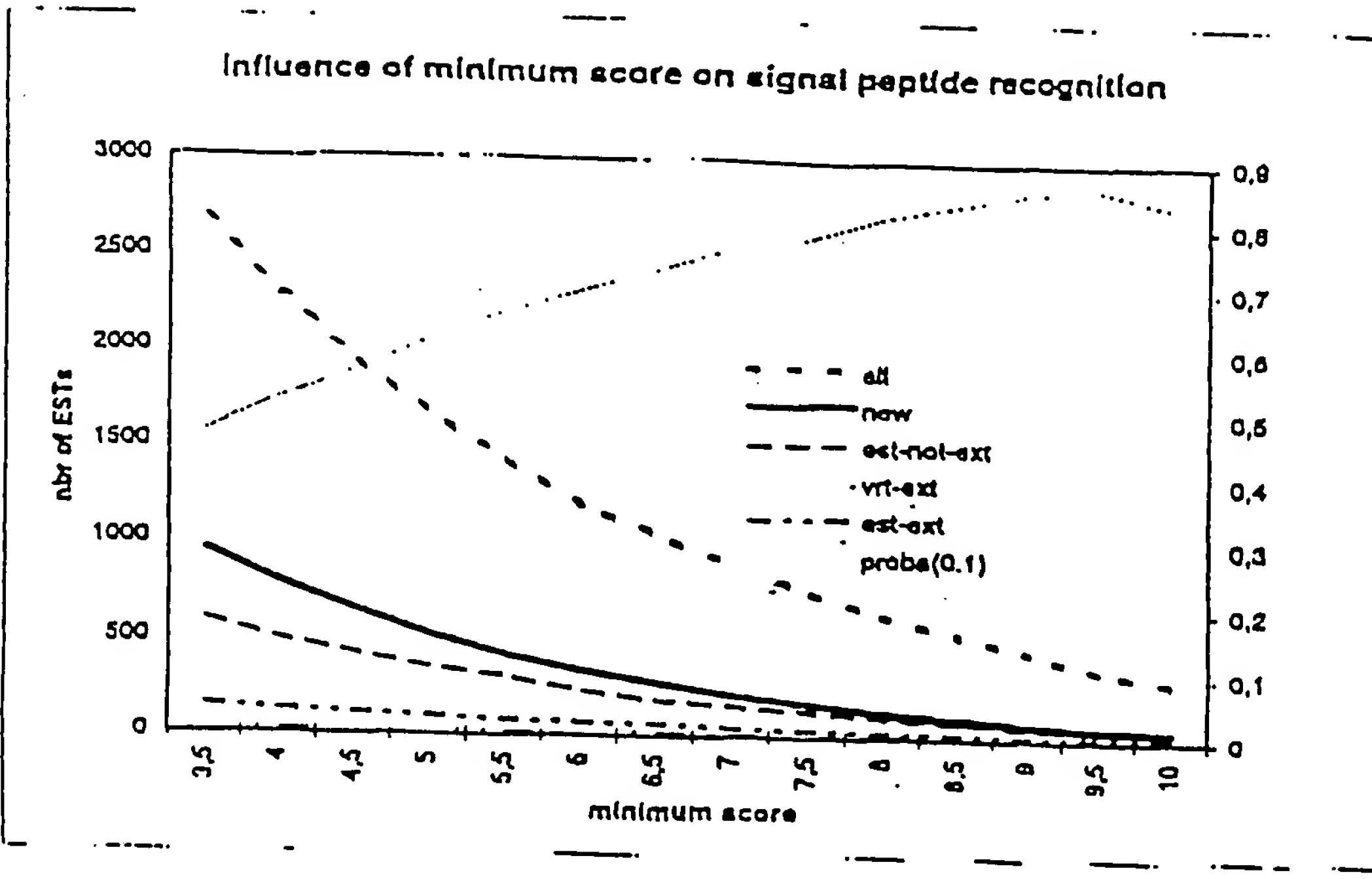


FIGURE 3

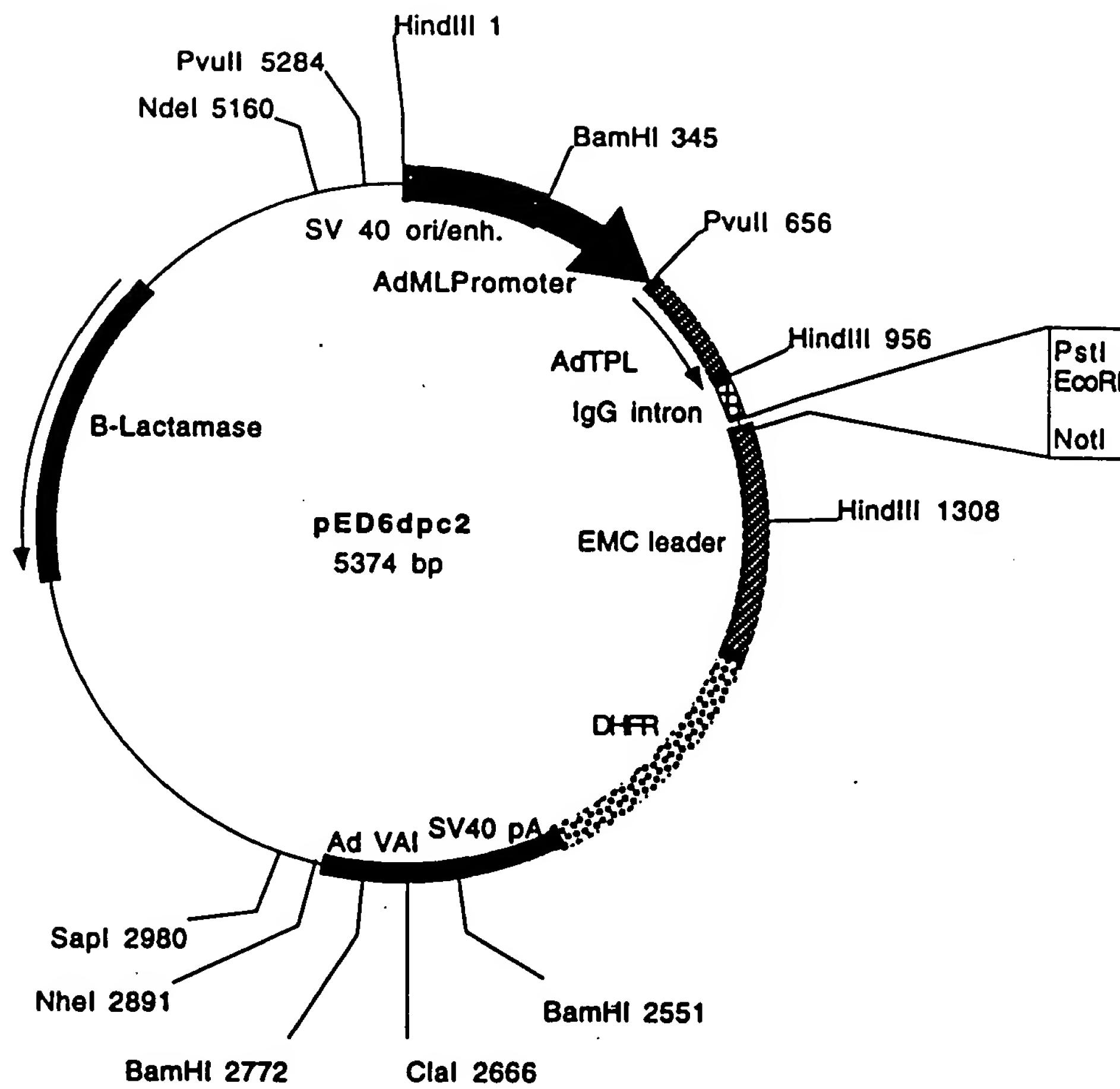
| Minimum signal peptide score | All ESTs | New ESTs | ESTs matching public EST closer than 40 bp from beginning | ESTs extending known mRNA more than 40 bp | ESTs extending public EST more than 40 bp |
|------------------------------|----------|----------|---|---|---|
| 3,5                          | 2874     | 947      | 599   | 23  | 150                                       |
| 4                            | 2278     | 784      | 488   | 23  | 126                                       |
| 4,5                          | 1943     | 647      | 425   | 22  | 112                                       |
| 5                            | 1657     | 523      | 353   | 21  | 86  |
| 5,5                          | 1417     | 418      | 307   | 19  | 80  |
| 6                            | 1180     | 340      | 238   | 18  | 68  |
| 6,5                          | 1035     | 280      | 186   | 18  | 60  |
| 7                            | 893      | 218      | 161   | 15  | 48  |
| 7,5                          | 753      | 173      | 132   | 12  | 36  |
| 8                            | 636      | 133      | 101   | 11  | 28  |
| 8,5                          | 543      | 104      | 83  | 8   | 26  |
| 9                            | 456      | 81       | 63  | 6   | 24  |
| 9,5                          | 364      | 57       | 48  | 6   | 18  |
| 10                           | 303      | 47       | 35  | 6   | 15  |

FIGURE 4

5/16

| Tissue                | All ESTs | New ESTs | ESTs matching public EST closer than 40 bp from beginning | ESTs extending known mRNA more than 40 bp | ESTs extending public EST more than 40 bp |
|-----------------------|----------|----------|---|---|---|
| Brain                 | 329      | 131      | 75  | 3   | 24  |
| Cancerous prostate    | 134      | 40       | 37  | 1   | 6   |
| Cerebellum            | 17       | 9        | 1   | 0   | 6   |
| Colon                 | 21       | 11       | 4   | 0   | 0   |
| Dystrophic muscle     | 41       | 18       | 8   | 0   | 1   |
| Fetal brain           | 70       | 37       | 18  | 0   | 1   |
| Fetal kidney          | 227      | 116      | 48  | 1   | 19  |
| Fetal liver           | 13       | 7        | 2   | 0   | 0   |
| Heart                 | 30       | 15       | 7   | 0   | 1   |
| Hypertrophic prostate | 86       | 23       | 22  | 2   | 2   |
| Kidney                | 10       | 7        | 3   | 0   | 0   |
| Large intestine       | 21       | 8        | 4   | 0   | 1   |
| Liver                 | 23       | 9        | 6   | 0   | 0   |
| Lung                  | 24       | 12       | 4   | 0   | 1   |
| Lung (cells)          | 57       | 38       | 6   | 0   | 4   |
| Lymph ganglia         | 163      | 60       | 23  | 2   | 12  |
| Lymphocytes           | 23       | 6        | 4   | 0   | 2   |
| Muscle                | 33       | 16       | 8   | 0   | 4   |
| Normal prostate       | 181      | 61       | 45  | 7   | 11  |
| Ovary                 | 90       | 57       | 12  | 1   | 2   |
| Pancreas              | 48       | 11       | 6   | 0   | 1   |
| Placenta              | 24       | 5        | 1   | 0   | 0   |
| Prostate              | 34       | 16       | 4   | 0   | 2   |
| Spleen                | 56       | 28       | 10  | 0   | 1   |
| Substantia nigra      | 108      | 47       | 27  | 1   | 6   |
| Surrenals             | 15       | 3        | 3   | 1   | 0   |
| Testis                | 131      | 68       | 25  | 1   | 8   |
| Thyroid               | 17       | 8        | 2   | 0   | 2   |
| Umbilical cord        | 65       | 17       | 12  | 1   | 3   |
| Uterus                | 28       | 15       | 3   | 0   | 2   |
| Non tissue-specific   | 568      | 48       | 177   | 2   | 28  |
| Total                 | 2677     | 947      | 601   | 23  | 150                                       |

FIGURE 5



Plasmid name: pED6dpc2

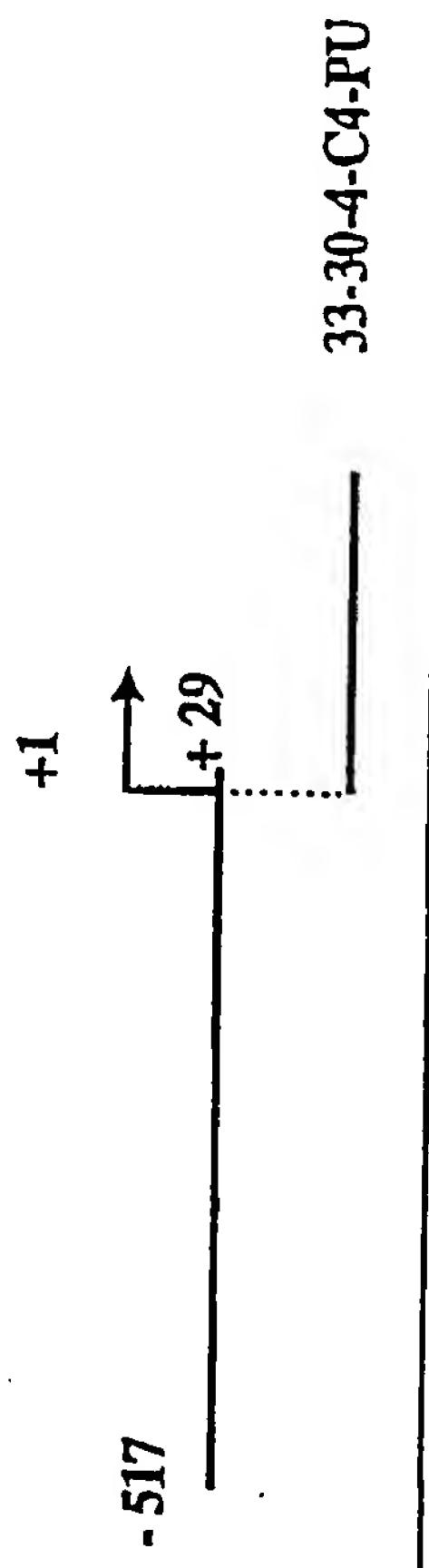
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

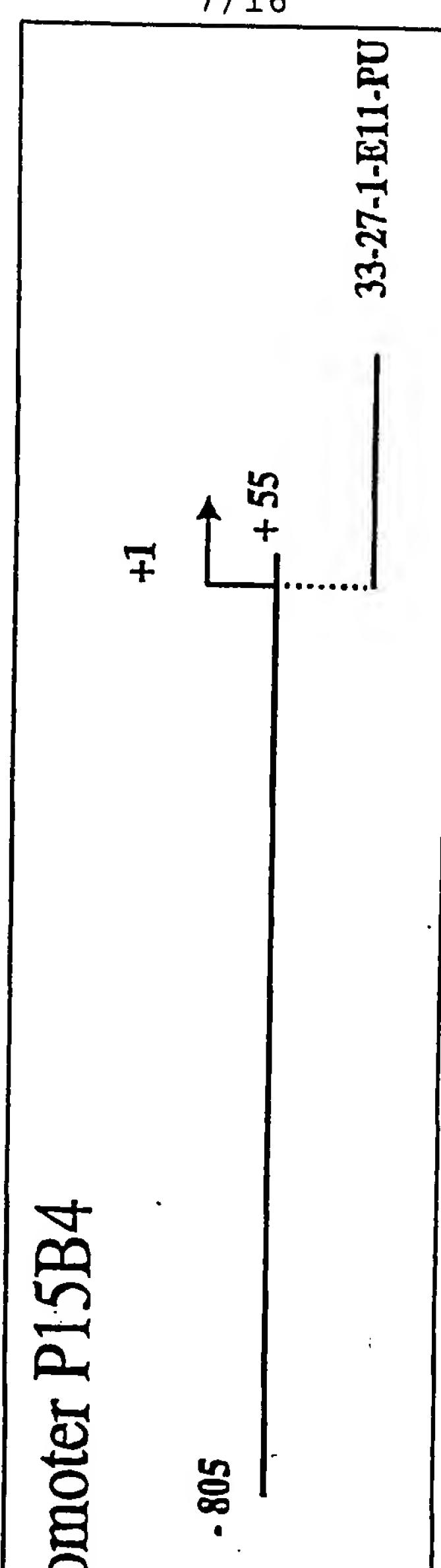
FIGURE 6

Description of Promoter structure isolated from SignalTag 5'ESTs

Promoter P13H2



Promoter P15B4



Promoter P29B6

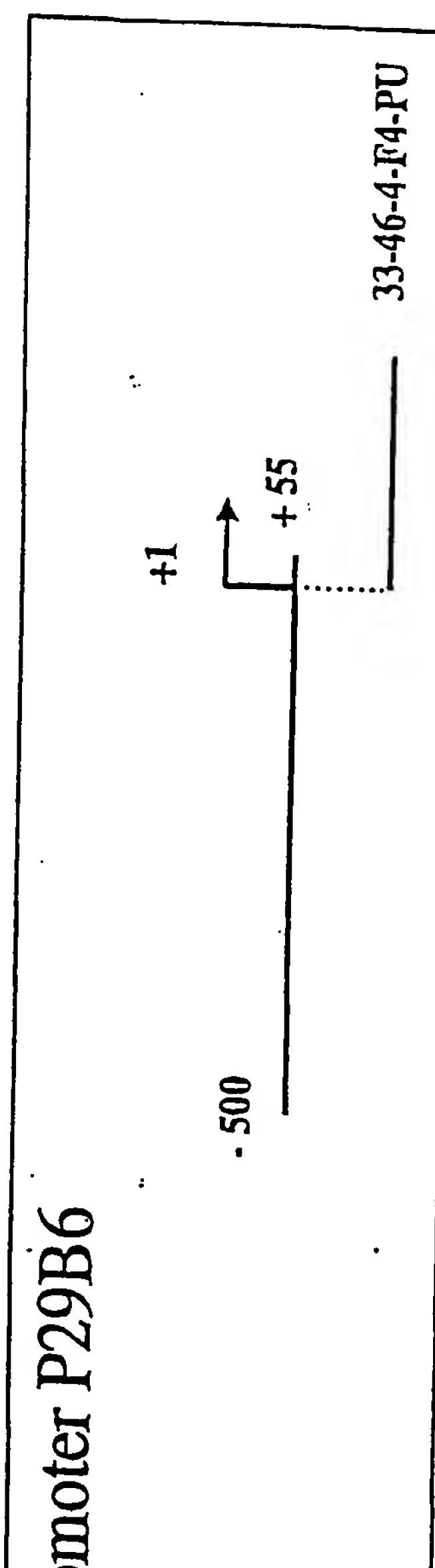


FIGURE 7

Description of Transcription Factor Binding Sites present on promoters isolated from SignalTag sequences.

Promoter sequence P13H2 (548 bp) :

| Matrix          | Position | Orientation | Score | Length | Sequence         |
|-----------------|----------|-------------|-------|--------|------------------|
| CMYB_01         | -602     | +           | 0.963 | 9      | TGTCAGTTG        |
| MYOD_Q6         | -501     | -           | 0.961 | 10     | CCCAACTGAC       |
| 88_01           | -444     | -           | 0.960 | 11     | AATAQAATTAG      |
| 88_01           | -425     | +           | 0.968 | 11     | AACTAAATTAG      |
| DELTAEF1_01     | -390     | -           | 0.960 | 11     | GCACACCTCAG      |
| GATA_C          | -364     | -           | 0.964 | 11     | AGATAAAATCCA     |
| CMYB_01         | -349     | +           | 0.958 | 9      | CTTCAGTTG        |
| GATA1_02        | -343     | +           | 0.959 | 14     | TTGTAGATAGGACA   |
| GATA_C          | -339     | +           | 0.953 | 11     | AGATAGGACAT      |
| TAL1ALPHAE47_01 | -235     | +           | 0.973 | 18     | CATAACAGATGGTAAG |
| TAL1BETAE47_01  | -235     | +           | 0.983 | 16     | CATAACAGATGGTAAG |
| TAL1BETATF2_01  | -235     | +           | 0.978 | 16     | CATAACAGATGGTAAG |
| MYOD_Q6         | -232     | -           | 0.954 | 10     | ACCATCTGTT       |
| GATA1_04        | -217     | -           | 0.953 | 13     | TCAAGATAAAGTA    |
| IK1_01          | -126     | -           | 0.963 | 18     | AGTTGGGAATTCC    |
| IK2_01          | -126     | -           | 0.985 | 12     | AGTTGGGAATTCC    |
| CREL_01         | -123     | -           | 0.962 | 10     | TGGGAATTCC       |
| GATA1_02        | -98      | -           | 0.960 | 14     | TCAGTGATATGGCA   |
| SRY_02          | -41      | -           | 0.951 | 12     | AAAACAAAAACA     |
| E2F_02          | -33      | +           | 0.957 | 8      | TTTAGCGC         |
| MZF1_01         | -6       | -           | 0.976 | 8      | TGAGGGGA         |

Promoter sequence P16B4 (861 bp) :

| Matrix      | Position | Orientation | Score | Length | Sequence     |
|-------------|----------|-------------|-------|--------|--------------|
| NFY_Q6      | -748     | -           | 0.958 | 11     | GGACCAATCAT  |
| MZF1_01     | -738     | +           | 0.962 | 8      | CCTGGGGA     |
| CMYB_01     | -684     | +           | 0.994 | 9      | TGACCGTTG    |
| VMYB_02     | -682     | -           | 0.985 | 9      | TCCAACGGT    |
| STAT_01     | -673     | +           | 0.988 | 9      | TTCTTGGAA    |
| STAT_01     | -673     | -           | 0.951 | 9      | TTCCAGGAA    |
| MZF1_01     | -658     | -           | 0.956 | 8      | TTGGGGGA     |
| IK2_01      | -451     | +           | 0.965 | 12     | GAATGGGATTC  |
| MZF1_01     | -424     | +           | 0.988 | 8      | AGAGGGGA     |
| SRY_02      | -398     | -           | 0.955 | 12     | QAAAACAAAACA |
| MZF1_01     | -216     | +           | 0.960 | 8      | QAAGGGGA     |
| MYOD_Q6     | -190     | +           | 0.981 | 10     | AGCATCTGCC   |
| DELTAEF1_01 | -176     | +           | 0.958 | 11     | TCCCACCTTC   |
| 88_01       | 6        | -           | 0.992 | 11     | GAGGCAATTAT  |
| MZF1_01     | 16       | -           | 0.986 | 8      | AGAGGGGA     |

Promoter sequence P29B6 (555 bp) :

| Matrix      | Position | Orientation | Score | Length | Sequence        |
|-------------|----------|-------------|-------|--------|-----------------|
| ARNT_01     | -311     | +           | 0.964 | 16     | GGACTCACGTCTACT |
| NMYC_01     | -309     | +           | 0.965 | 12     | ACTCACGTGCTG    |
| USF_01      | -309     | +           | 0.985 | 12     | ACTCACGTGCTG    |
| USF_01      | -309     | -           | 0.985 | 12     | OAGCAOGTGAQT    |
| NMYC_01     | -309     | -           | 0.956 | 12     | OAGCACGTGAQT    |
| MYCHMAX_02  | -309     | -           | 0.972 | 12     | CAGCACGTGAQT    |
| USF_C       | -307     | +           | 0.997 | 8      | TCACGTGTC       |
| USF_C       | -307     | -           | 0.991 | 8      | GCACGTGA        |
| MZF1_01     | -292     | -           | 0.988 | 8      | CATGGGGA        |
| ELK1_02     | -105     | +           | 0.963 | 14     | CTCTCCGQAAGCCT  |
| CET81P54_01 | -102     | +           | 0.974 | 10     | TCGGGAAGCC      |
| AP1_Q4      | -42      | -           | 0.963 | 11     | AGTGAATGAAAC    |
| AP1FJ_Q2    | -42      | -           | 0.981 | 11     | AGTGAATGAAAC    |
| PADS_C      | 45       | +           | 1.000 | 9      | TGTGGTCTC       |

Figure 8

9 / 16

98.2% identity in 113 aa overlap

|          |  |  |    |     |     |    |
|----------|--|--|----|-----|-----|----|
| 10       | 20   | 30   | 40 | 50  | 60  |    |
| SeqID214 | MVIRVYIASSSGSTA<br>:::                     | KKKQQDVLF<br>LEANKIGFEEKDIA<br>AANNEENRKWMREN<br>VPENSRA |    |     |     |    |
| AF042081 | MVIRVYIASSSGSTA<br>:::                     | KKKQQDVLF<br>LEANKIGFEEKDIA<br>AANNEENRKWMREN<br>VPENSRA |    |     |     |    |
|          | 10   | 20   | 30 | 40  | 50  | 60 |
|          |  |  |    |     |     |    |
|          | 70   | 80   | 90 | 100 | 110 |    |
| SeqID214 | TGN <u>PLPPQIFN</u> ESQYRGDYDAFFEAR<br>::: | ENNNAVYAFLGLTAPSGS<br>KEAEVQAKQQ                         |    |     |     |    |
| AF042081 | TGY <u>PLPPQIFN</u> ESQYRGDYDAFFEAR<br>::: | ENNNAVYAFLGLTAPP<br>GSKEAEVQAKQQ                         |    |     |     |    |
|          | 70   | 80   | 90 | 100 | 110 |    |

FIGURE 9

10/16

|          |  |
|----------|--|
| seqID215 | MADDLKRFLYKKLPSVEGLHAIIVVSDRGVPVIKVANDNAPEHALRPGFLSTFALATDQG           |
| seqID185 | MADDLKRFLYKKLPSVEGLHAIIVVSDRGVPVKVANDNAPEHALRPGFLSTFALATDQG            |
| AF082526 | MADDLKRFLYKKLPSVEGLHAIIVVSDRGVPVIKVANDSAPEHALRPGFLSTFALATDQG<br>*****  |
| seqID215 | SKLGLSKNKSIIICYYNTYQVVQFNRLPLVVSFIASSSANTGLIVSLEKELAPLFEELRQV          |
| seqID185 | SKLGLSKNKSIIICYYNTYQVVQFNRLPLVVSFIASSSANTGLIVSLEKELAPLFEELRQV          |
| AF082526 | SKLGLSKNKSIIICYYNTYQVVQFNRLPLVVSFIASSSANTGLIVSLEKELAPLFEELIKV<br>***** |
| seqID215 | VEVS   |
| seqID185 | VEVS   |
| AF082526 | VEVS<br>****   |

FIGURE 10

11/16

91.3% identity in 230 aa overlap

|   |     |     |     |     |     |
|---|-----|-----|-----|-----|-----|
| 10  | 20  | 30  | 40  | 50  | 60  |
| SeqID186 MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTG<br>:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::  |     |     |     |     |     |
| AF072128 MASLGVLVGYILGLLGLLGTASIAMLLPNWRTSSYVGASIVTAVGFSKGLWMECATHSTG<br>10 20 30 40 50 60  |     |     |     |     |     |
| 70  | 80  | 90  | 100 | 110 | 120 |
| SeqID186 ITQCDIYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVA<br>:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::. |     |     |     |     |     |
| AF072128 ITQCDIYSTLLGLPADIQAAQAMMVTSSAMSSLACIISVVGMRCTVFCQDSRAKDRVAVV<br>70 80 90 100 110 120   |     |     |     |     |     |
| 130   | 140 | 150 | 160 | 170 | 180 |
| SeqID186 GGVFFILGGLLGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGII<br>:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::. |     |     |     |     |     |
| AF072128 GGVFFILGGILGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISALFSLVAGVI<br>130 140 150 160 170 180  |     |     |     |     |     |
| 190   | 200 | 210 | 220 | 230 |     |
| SeqID186 LCFSCSSQRNRSNYYDAYQAQPLATRSSPRPGQPPKVSEFNSYSLTGYV<br>:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.....:::.                    |     |     |     |     |     |
| AF072128 LCFSCSPQGNRTNYYDGYQAQPLATRSSPRSAQQPKAKSEFNSYSLTGYV<br>190 200 210 220 230  |     |     |     |     |     |

FIGURE 11

12/16

98.3% identity in 121 aa overlap

|          |   |  |
|----------|---|--|
| seqID213 | 10            20            30  |  |
|          | RFRKETDNAIIMKVDKDRQMVVLEEEFRNISPEELKME                                      |  |
|          | ::::::: ::::::::::::::::::::: :::::::::::::                                 |  |
| AB001993 | MSDSLVVCEVDPELTEKLKRFRKETDNAIIMKVDKDRQMVVLEEEFQNIISPEELKME                  |  |
|          | 10            20            30            40            50            60    |  |
|          | 40            50            60            70            80            90    |  |
| seqID213 | LPERQPRFVVYSYKYVRDDGRVSYPLCFIFSSPVGCKPEQQMMYAGSKNRLVQTAELTKV                |  |
|          | ::::::: ::::::::::::::::::::: :::::::::::::                                 |  |
| AB001993 | LPERQPRFVVYSYKYVHDDGRVSYPLCFIFSSPVGCKPEQQMMYAGSKNRLVQTAELTKV                |  |
|          | 70            80            90            100            110            120 |  |
|          | 100            110            120   |  |
| seqID213 | FEIRTTDDLTEAWLQEKLSSFR  |  |
|          | ::::::: :::::::::::::   |  |
| AB001993 | FEIRTTDDLTEAWLQEKLSSFR  |  |
|          | 130            140  |  |

FIGURE 12

95.6% identity in 91 aa overlap

|           |  |    |                       |       |     |     |    |
|-----------|--|----|-----------------------|-------|-----|-----|----|
| seq ID191 |  |    | 10                    | 20    |     |     |    |
|           |  |    | MGCVFQSTEDKCIFKIDWTLS |       |     |     |    |
| W36955    | MFCPLKLILLPVLLDYSLGLNDLNVSPP<br>ELTVHVGDSALMGCVFQSTEDKCIFKIDWTLS |    | .....                 | ..... |     |     |    |
|           | 10   | 20 | 30                    | 40    | 50  | 60  |    |
| seq ID191 | PGEHAKDEYVLYYSNL<br>SVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRL  | 30 | 40                    | 50    | 60  | 70  | 80 |
| W36955    | PGEHAKDEYVLYYSNL<br>SVPIGRFQNRVHLMGDNL<br>CNDGSLLLQDVQDV<br>E    | 70 | 80                    | 90    | 100 | 110 |    |
| seq ID191 | KGESQVFKKAVVLHVLPEEPKG<br>TQMLT                                  | 90 | 100                   |       |     |     |    |

FIGURE 13

99.0% identity in 381 aa overlap;

|          | 10  | 20 | 30 | 40 | 50 | 60 |
|----------|---|----|----|----|----|----|
| seqID200 | MLLSIGMLMLSATQVYTQLFAFLNPLPVEADILAYNFENASQTFDDLPARFGYRLP<br>::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::     |    |    |    |    |    |
| AF037204 | MLLSIGMLMLSATQVYTILTVQLFAFLNLLPVEADILAYNFENASQTFDDLPARFGYRLP<br>10 20 30 40 50 60   |    |    |    |    |    |
|          | 70 80 90 100 110 120  |    |    |    |    |    |
| id200    | AEGLKGFLINSKPENACEPIVPPPVDNSSGTFIVLIRRLDCNFDIKVLNAQRAGYKAAI<br>::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::  |    |    |    |    |    |
| AF037204 | AEGLKGFLINSKPENACEPIVPPPVDNSSGTFIVLIRRLDCNFDIKVLNAQRAGYKAAI<br>70 80 90 100 110 120   |    |    |    |    |    |
|          | 130 140 150 160 170 180   |    |    |    |    |    |
| id200    | VHNVDSDLISMGSNDIEVLKKIDIPSVFIGESSASSLKDEFTYEKGHLILVPEFSLPL<br>::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::   |    |    |    |    |    |
| AF037204 | VHNVDSDLISMGSNDIEVLKKIDIPSVFIGESSANSLKDEFTYEKGHLILVPEFSLPL<br>130 140 150 160 170 180   |    |    |    |    |    |
|          | 190 200 210 220 230 240   |    |    |    |    |    |
| id200    | EYYLIPFLIIVGICLILIVIFMITKFVQDRHRARRNRLRDKQLKKLPVHKFKKGDEYDVC<br>::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: |    |    |    |    |    |
| AF037204 | EYYLIPFLIIVGICLILIVIFMITKFVQDRHRARRNRLRDKQLKKLPVHKFKKGDEYDVC<br>190 200 210 220 230 240   |    |    |    |    |    |
|          | 250 260 270 280 290 300   |    |    |    |    |    |
| id200    | AICLDEYEDGDKLRILPCSHAYHCKCVDPWLTKTKKTCPVCRQKVVPSEQDSDSDTDSSQ<br>::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: |    |    |    |    |    |
| AF037204 | AICLDEYEDGDKLRILPCSHAYHCKCVDPWLTKTKKTCPVCKQKVVPSEQDSDSDTDSSQ<br>250 260 270 280 290 300   |    |    |    |    |    |
|          | 310 320 330 340 350 360   |    |    |    |    |    |
| id200    | EENEVTEHTPLLRPLASVSAQSGALSESRSHQNMTESSDYEEDDNEDTSSDAENEINE<br>::::::::::: ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::   |    |    |    |    |    |
| AF037204 | EENEVTEHTPLLRPLASVSAQSGALSESRSHQNMTESSDYEEDDNEDTSSDAENEINE<br>310 320 330 340 350 360   |    |    |    |    |    |
|          | 370 380   |    |    |    |    |    |
| id200    | HDVVVQLQPNGERDYNIAINTV<br>::::::::::: :::::::::::   |    |    |    |    |    |
| AF037204 | HDVVVQLQPNGERDYNIAINTV<br>370 380   |    |    |    |    |    |

### FIGURE 14

100.0% identity in 68 aa overlap

|          |   |       |    |    |    |    |
|----------|---|-------|----|----|----|----|
|          | 10  | 20    | 30 | 40 | 50 | 60 |
| seqID192 | MSVFWGVGFLVPWFIPKGPNRGVIITMLVTCSVCCYLFWLIAILAQLNPLFGPQLKNET | :     | :  | :  | :  | :  |
| Y15286   | MSVFWGVGFLVPWFIPKGPNRGVIITMLVTCSVCCYLFWLIAILAQLNPLFGPQLKNET | 20    | 30 | 40 | 50 | 60 |
|          |   |       |    |    |    | 70 |
| seqID192 | IWYLYKYHW   |       |    |    |    |    |
|          | :   | ::::: |    |    |    |    |
| Y15286   | IWYLYKYHW   |       |    |    |    |    |
|          |   | 80    |    |    |    |    |

FIGURE 15

|                                |  |
|--------------------------------|--|
| seqID201<br>seqID227<br>X85116 | -MDSRVS--SPEKQDKENFVGNNKRLGVCGLWILFSLSFLVIITFPISIWMCLKIIREY<br>-----MWLDP-----VFPLFPVG-----DH<br>MAEKRHTRDSEAQR LPDSFKDSPSKGLGPCGWILVAFSFLFTVITFPISIWMCIKIIKEY<br>* . * * .  |
| seqID201<br>seqID227<br>X85116 | ERAVVFR LGRIQADKAKGPGLILVLP CIDVFVKVDLRTVTCNIPPPQEILTRDSVTTQVDG<br>Y-----LPHLHMDVLEG--LILVLP CIDVFVKVDLRTVTCNIPPPQEILTRDSVTTQVDG<br>ERAIIIFRLGRILQGGAKGPGLFFILPCTDSFIKVDMRTISFDIPPQEILTKDSVTISVDG<br>* .. * * . * * . * * . * * . * * . * * * * * * . * * * * * * . * * * *          |
| seqID201<br>seqID227<br>X85116 | VYYYRIYSAVSAVANVNDVHQATFLLAQTTLRNVLTQTLSQILAGREEIAHSIQTLDD<br>VYYYRIYSAVSAVANVNDVHQATFLLAQTTLRNVLTQTLSQILAGREEIAHSIQTLDD<br>VYYRVQNATLAVANITNADSATRLLAQTTLRNVLTQTLSQILSDREEIAHNMQSTLDD<br>* * * . * * * . . * * * * * * * * * . * * * . * * * * * * . * * * * . * * * *              |
| seqID201<br>seqID227<br>X85116 | ATELWGIRVARVEIKDV RIPVQLQRSMAAEAEATREARAKVLAEGEMSASKSLKSASMV<br>ATELWGIRVARVEIKDV RIPVQLQRSMAAEAEATREARAKVLAEGEMNASKSLKSASMV<br>ATDAWGIKVERVEIKDV KLPVQLQRAMAAEAEASREARAKVIAAEGEMNASRALKEASMV<br>* * . * * * . * * * * * . * * * * * . * * * * * . * * * * * . * * * * . * * . * * * |
| seqID201<br>seqID227<br>X85116 | LAESPIALQLRYLQTLSTVATEKNSTIVFPLPMNILEGIGGVSYDNHKKLPNKA<br>LAESPIALQLRYLQTLSTVATEKNSTIVFPLPMNILEGIGGVSYDNHKKLPNKA<br>ITESPAALQLRYLQTLTTIAAEKNSTIVFPLPIDMLQGIIGAKHSHLG-----<br>* * * * * * * * * * . * * * * * * * * * * . * * * * * . * * * *   |

## FIGURE 16